

A Discrepancy based Approach to Integer Programming

Karthekeyan Chandrasekaran*

Santosh Vempala†

Abstract

We consider integer programs on polytopes in \mathbb{R}^n with m facets whose normal vectors are chosen independently from any spherically symmetric distribution. We show that for m at most $2\sqrt{n}$, there exist constants $c_1 < c_2$ such that with high probability, this random IP is infeasible if the largest ball contained in the corresponding polytope has radius less than $c_1\sqrt{\log(2m/n)}$ and it is feasible if the radius is at least $c_2\sqrt{\log(2m/n)}$. Thus, a transition from infeasibility to feasibility happens within a constant factor increase in the radius. Moreover, if the polytope contains a ball of radius $\Omega(\log(2m/n))$, then there is a randomized polynomial-time algorithm to find an integer solution with high probability (over the input). Our main tools are: a new connection between integer programming and matrix discrepancy, a bound on the discrepancy of random Gaussian matrices and Bansal's algorithm for finding low-discrepancy solutions.

*karthe@gatech.edu, Georgia Institute of Technology

†vempala@cc.gatech.edu, Georgia Institute of Technology

1 Introduction

Integer Linear Programming (IP) is a general and powerful formulation [19] widely used to address combinatorial problems. One standard variant is the integer feasibility problem: given a polytope P specified by linear constraints $Ax \leq b$, find an integer solution in P or report that none exists. The problem is NP-hard and appears in Karp’s original list [12]. Dantzig [6] suggested the possibility of IP being a *complete* problem even before the Cook-Levin theory of NP-completeness.

Even though integer programs are NP-hard, they are solved routinely in practice. Most LP solvers in the market have an IP solver that employ heuristics based on cutting plane and branch and bound techniques to solve the IP. In spite of a vast and rich literature on the theory of cutting planes, there are few complexity guarantees apart from an exponential bound due to Gomory for binary integer programs [9]. The best-known rigorous bound on the complexity of general IP is $n^{O(n)}$ [11].

Some special cases of IP can be solved efficiently, e.g., IP instances where the constraint matrix A specifying the polytope is totally unimodular; more generally, when the integer hull of the polytope has an efficient separation oracle (using the ellipsoid algorithm [10]). Motivated by the question, “*which IP’s can be solved in polynomial time?*”, we consider probabilistic integer programs, as defined next.

Model. A random integer program $P = P(n, m, x_0, R)$ is generated as follows: we pick a random $m \times n$ matrix A with i.i.d. rows from a spherically symmetric distribution; and a vector b such that the hyperplane for each constraint is at distance at least R from x_0 , i.e., $b_i \geq R\|A_i\| + A_i x_0$, where A_i is the i ’th row of A .

The condition above implies that P contains a ball of radius R centered at x_0 . We study the feasibility of P as a function of the radius R . Our existence bound is independent of the center x_0 and our algorithm for finding a feasible solution is only given A, b .

1.1 Results

We are able to prove the following bounds on the feasibility of random integer programs.

Theorem 1. *Let $m \geq 1000n$ and*

$$R_0 = \sqrt{\frac{1}{6} \log \frac{2m}{n}}, \quad R_1 = 128 \left(\sqrt{\log \frac{2m}{n}} + \sqrt{\frac{\log m \log(mn) \log(2m/\log m)}{n}} \right).$$

Then, with probability at least $1 - 2me^{-n/96}$,

1. *for every $x_0 \in \mathbb{R}^n$, the random polytope $P(n, m, x_0, R_1)$ contains an integer point, and*
2. *for $x_0 = (1/2, \dots, 1/2)$, the random polytope $P(n, m, x_0, R_0)$ does not contain an integer point.*

We note that for $m = 2^{O(\sqrt{n})}$, the second term in R_1 is of the same order as the first and so the two thresholds are within a constant factor of each other. When $m = O(n)$, the transition between infeasibility and feasibility happens between two absolute constants.

The upper bound in Theorem 1 is based on showing that given m vectors $A_1, \dots, A_m \in \mathbb{R}^n$ such that each A_i is a random unit vector, and a point $x_0 \in \mathbb{R}^n$, there exists a point $x \in \mathbb{Z}^n$

satisfying $|A_i(x - x_0)| \leq R_1$ for every $i \in [m]$. To prove this, we consider the discrepancy of a random Gaussian matrix.

Theorem 2. *Let $A \in \mathbb{R}^{m \times n}$ be a random matrix with i.i.d. entries from $N(0, \sigma^2)$. For any $x_0 \in \mathbb{R}^n$, with high probability, there exists a point $x \in \mathbb{Z}^n$ obtained by rounding each coordinate of x_0 either up or down such that, for every $i \in [m]$,*

$$|A_i(x - x_0)| \leq \frac{1}{2} \sigma \sqrt{n} R_1.$$

The upper bound in Theorem 1 follows from Theorem 2, by choosing $\sigma^2 = 1/n$ and observing that m random unit vectors are obtained by scaling each row by at most a constant with probability at least $1 - 2me^{-n/96}$. In terms of classical discrepancy theory, Theorem 2 is equivalent to a bound of $(\sigma R_1 \sqrt{n}/2)$ on the linear discrepancy of a random Gaussian matrix.

Algorithm to find an integer point. We complement our existence theorem with an algorithm to find an integer point. With $R = \Omega(\sqrt{\log m})$ and $x_0 = (1/2, \dots, 1/2)$, there is a trivial algorithm — pick a random 0/1 vector. Most such vectors will be feasible in $P(n, m, x_0, R)$. But with smaller R , and arbitrary centers x_0 , only an exponentially small fraction of nearby integer vectors might be feasible, so such direct sampling/enumeration would not give a feasible integer point.

Theorem 3. *Given a random IP $P = P(n, m, x_0, R)$ where*

$$R \geq 2^{16} \left(\log \frac{2m}{n} + \sqrt{\frac{\log m \log(mn)}{n}} \log \frac{2m}{\log m} \right)$$

there is a randomized polynomial-time algorithm to find an integer point $x \in P$ with probability at least $1 - 2me^{-n/96}$.

Our algorithm builds on Bansal's semidefinite programming based method for finding low-discrepancy solutions [1]. His algorithm finds a vector $X \in \{-1, +1\}^n$ with discrepancy $O(\sqrt{n} \log(2m/n))$ when A is any 0/1 matrix. We adapt his algorithm to find a vector $x \in \{0, 1\}^n$ such that $\|A(x - x_0)\|_\infty$ is $O(\sigma \sqrt{n} \log(2m/n))$ when each entry in A is from $N(0, \sigma^2)$. We note that the resulting algorithm can be viewed as a general rounding method for integer programming; we are able to analyze it only for random IPs.

1.2 The connection to discrepancy

The main conceptual contribution of our paper is a connection between integer feasibility and discrepancy theory [15, 20, 21].

Suppose we seek $-1/1$ points in the polytope (as opposed to integer points). Given a matrix $A \in \mathbb{R}^{m \times n}$, consider the polytope $P(A, r) = \{x \in \mathbb{R}^n : |A_i x| \leq r \ \forall i \in [m]\}$ for a fixed positive r . The discrepancy of a matrix A is defined to be the least r so that the polytope $P(A, r)$ contains a $-1/1$ point. More formally,

$$\text{disc}(A) := \min_{x \in \{-1, +1\}^n} \|Ax\|_\infty$$

Thus, if we can evaluate the discrepancy of the constraint matrix A describing a polytope, then we can obtain a characterization of $-1/1$ -feasibility of the polytope.

The following related notion of linear discrepancy helps in characterizing integer feasibility of the polytope:

$$\text{lindisc}(A) := \max_{x_0 \in [0,1]^n} \min_{x \in \{0,1\}^n} \|A(x - x_0)\|_\infty.$$

We observe that every polytope $P = \{x \in \mathbb{R}^n \mid |A_i(x - x_0)| \leq b_i \text{ for } i \in [m]\}$ where $b_i \geq \text{lindisc}(A)$ contains an integer point for every $X_0 \in \mathbb{R}^n$. This is because, by linear transformation, we may assume that x_0 is in the fundamental cube defined by the standard basis unit vectors. Thus, if each row of the constraint matrix is a unit vector, then linear discrepancy of the constraint matrix gives one possible radius of the largest inscribed ball so that the polytope is integer feasible for every center x_0 .

This approach to verify integer feasibility of a polytope fails for arbitrary polytopes since it is NP-hard to find the discrepancy of a set-system to within a factor of \sqrt{n} [5]. We show that this approach can still be used for random polytopes due to tight bounds on the discrepancy and the linear discrepancy of random matrices.

The central quantity that leads to all known bounds on discrepancy and linear discrepancy in the literature is hereditary discrepancy defined as follows:

$$\text{herdisc}(A) := \max_{S \subseteq [n]} \text{disc}(A^S)$$

where A^S denotes the submatrix of A containing columns indexed by the set S . The best known bound on discrepancy and hereditary discrepancy of arbitrary matrices is due to Spencer [20].

Theorem 4. [20] *For any matrix $A \in \mathbb{R}^{m \times n}$ and any subset $S \subseteq [n]$, there exists a point $z \in \{-1, +1\}^{|S|}$ such that*

$$|A_i^S z| \leq 11 \sqrt{|S| \log \frac{2m}{|S|}} \max_{i \in [m], j \in S} |A_{ij}|.$$

for every $i \in [m]$.

Lovász, Spencer and Vesztergombi [14] showed the following relation between hereditary discrepancy and linear discrepancy.

Theorem 5. [14] *For any matrix A , $\text{lindisc}(A) \leq \text{herdisc}(A)$.*

Hence, every polytope $P = \{x \in \mathbb{R}^n \mid |A_i(x - x_0)| \leq b_i \text{ for } i \in [m]\}$ where

$$b_i = \Omega\left(\max_{i \in [m], j \in [n]} |A_{ij}| \sqrt{n \log(2m/n)}\right)$$

contains an integer point for every $X_0 \in \mathbb{R}^n$.

In our setting, each entry A_{ij} is from $N(0, \sigma^2)$. Using standard concentration for $|A_{ij}|$ and a union bound to bound the maximum entry $|A_{ij}|$ leads to the following weak bound whp: every polytope $P = \{x \in \mathbb{R}^n \mid |A_i(x - x_0)| \leq b_i \text{ for } i \in [m]\}$ with $b_i = \Omega(\sigma \sqrt{n \log mn \log(2m/n)})$ contains an integer point for any $x_0 \in \mathbb{R}^n$. We strengthen this result using a nonstandard normalization in the proof of Spencer's result.

Our infeasibility threshold is also based on discrepancy. We begin with a lower bound on linear discrepancy of random matrices, which excludes any 0/1 vector from being a solution for the chosen radius of the inscribed ball, then extend this to exclude all integer points.

1.3 Related work in IP

Probabilistic instances of several combinatorial problems have been studied and shown to have efficient algorithms, e.g., random knapsack [2], feedback vertex set, largest clique, chromatic number, etc in random graphs [4]. These algorithms are combinatorial in nature. To our knowledge, the first work on probabilistic instances of IP was by Furst and Kannan [8]. They consider a probabilistic instance of the subset-sum problem. The subset-sum problem is to find $x \in \{0, 1\}^n$ satisfying $Ax = b$ where $A \in \mathbb{N}^n$, $b \in \mathbb{Z}$ or output a short certificate of infeasibility. In the probabilistic instance that they consider, the coefficients of A are drawn uniformly at random from the discrete set $\{1, 2, \dots, M\}$ for some large M . For $M \geq 2^{n^2/2+2n}n^{3n/2}$, they show that there exists a polynomial time deterministic algorithm to solve the subset-sum problem with high probability for any b .

Pataki et al. [17] generalize this further to address probabilistic integer programming problems with m constraints where each entry in the constraint matrix is drawn i.i.d. from $\{1, \dots, M\}$. They consider the integer feasibility problem of the polytope defined by the following linear constraints:

$$\begin{aligned} l_1 &\leq AX \leq w_1 \\ l_2 &\leq X \leq w_2 \end{aligned}$$

where $A \in \{1, 2, \dots, M\}^{m \times n}$, $l_1, w_1 \in \mathbb{R}^m$, $l_2, w_2 \in \mathbb{R}^n$. They show that if each entry in the constraint matrix A is drawn uniformly at random from the discrete set $\{1, 2, \dots, M\}$ where $M \geq (2^{(n+4)/2} \|(w_1; w_2) - (l_1; l_2)\|)^{n/(m+1)}$, then there exists a polynomial time deterministic algorithm that solves the integer feasibility problem with high probability.

In other related work, Beier and Vöcking [3] and Röglin and Vöcking [18] give a smoothed analysis for some special cases of IP.

2 Preliminaries

We will use the following standard tail bounds.

Lemma 1. *Let Y be a random variable distributed according to $N(0, \sigma^2)$. Then for any $t > 0$,*

$$\Pr(|Y| \leq t\sigma) \leq \min \left\{ 1 - \sqrt{\frac{2}{\pi}} \left(\frac{t}{t^2 + 1} \right) e^{-\frac{t^2}{2}}, t\sqrt{\frac{2}{\pi}} \right\}.$$

Lemma 2. *If X is drawn from the Gaussian distribution $N(0, \sigma^2)$, then for any $\lambda \geq 1$*

$$\Pr(|X| \geq \lambda\sigma) \leq 2e^{-\frac{\lambda^2}{2}}.$$

Lemma 3. [1] *Let $0 = X_0 = X_1, \dots, X_n$ be a martingale with increments $Y_i = X_i - X_{i-1}$. Suppose for $1 \leq i \leq n$, we have that $Y_i | (X_{i-1}, \dots, X_0)$ is distributed as Gaussian $N(0, \eta_i^2)$ where η_i is a constant such that $|\eta_i| \leq 1$. Then*

$$\Pr(|X_n| \geq \lambda\sqrt{n}) \leq 2e^{-\lambda^2/2}.$$

Lemma 4. [7] *If random variables X_1, \dots, X_r are drawn i.i.d. from the normal distribution $N(0, \sigma^2)$, then for any $\lambda > 0$*

$$\Pr \left(\left| \sum_{j \in [r]} X_j^2 - r\sigma^2 \right| \geq \lambda\sqrt{r}\sigma^2 \right) \leq 2e^{-\frac{\lambda^2}{24}}.$$

Lemma 5. *For any subset $S \subseteq [n]$ and for any fixed set of vectors $a_i, i \in [m]$, if each coordinate $X_j, j \in [n]$ is drawn uniformly at random from the set $\{-1, +1\}$, then*

$$\Pr \left(\left| \sum_{j \in S} a_{ij} X_j \right| \geq \lambda \right) \leq 2e^{-\frac{\lambda^2}{2 \sum_{j \in S} a_{ij}^2}}.$$

For a matrix $A \in \mathbb{R}^{m \times n}$ and any $S \subseteq [n]$, let A_i^S denote the vector A_i restricted to the coordinates in S . We say that the discrepancy of a vector A_i due to x is at most t_i if $|A_i x| \leq t_i$. We say that x incurs a discrepancy of at most t if $\max_{i \in [m]} |A_i x| \leq t$.

3 Linear discrepancy of a random matrix

By Theorem 5, it is sufficient to bound the hereditary discrepancy of random matrix to obtain an upper bound on the linear discrepancy of random matrix. We show the following bound on hereditary discrepancy of random matrix.

Theorem 6. *Suppose we have m vectors $A_1, \dots, A_m \in \mathbb{R}^n$, such that A_{ij} is drawn from the distribution $N(0, \sigma^2)$ for each $i \in [m], j \in [n]$. Then, for any $S \subseteq [n]$, with high probability, there exists a point $x \in \{-1, 1\}^{|S|}$ such that,*

$$|A_i^S x| \leq 32\sigma \left(\sqrt{n \log \frac{2m}{n}} + \sqrt{\log m \log mn \log \frac{2m}{\log m}} \right) \text{ for every } i \in [m].$$

Remark: The bound on the discrepancy of submatrix A^S given in Theorem 6 is independent of the size of S . This is unlike Spencer's result (Theorem 4) where the discrepancy of A^S is bounded by a function of $|S|$.

Our overall strategy is similar to that of Spencer (and subsequent work): We first show that there exists a point $z \in \{0, -1, +1\}^{|S|}$ with at least $|S|/2$ non-zero coordinates such that $|A_i^S z|$ is small. We start with $x = 0, S = [n]$ and use z to fix at least half of the coordinates of x to $+1$ or -1 . Then we take S to be the set of coordinates that are set to zero in the current x and use z to fix at least half of the remaining coordinates of x to $+1$ or -1 . We repeat this until all coordinates of x are non-zero. Since at most $|S|/2$ coordinates are set to zero in each round of fixing coordinates, we will repeat at most $\log n$ times. The total discrepancy is bounded by the sum of the discrepancies incurred in each round of fixing. The discrepancy incurred by fixing those coordinates of x which are non-zeros in z is bounded by

$$|A_i^S x^S| = |A_i^S z| \leq 4 \|A_i^S\| \sqrt{\log \frac{2m}{|S|}}$$

for all vectors $A_i, i \in [m]$ and all subsets $S \subseteq [n]$.

This general bound depends on the length of the vector A_i^S . It is straightforward to obtain $\|A_i^S\| \leq 2\sigma \sqrt{|S| \log mn}$ whp in our setting by bounding the maximum coefficient. This gives a bound of

$$8\sigma \sqrt{|S| \log(mn) \log \frac{2m}{|S|}}$$

on the discrepancy of A^S , i.e., A restricted to any subset S of columns. This bound on the discrepancy of A^S is good enough when the cardinality of S is smaller than some threshold, but too large for large S . E.g., when $S = [n]$, this gives a total discrepancy of at most $O(\sigma\sqrt{n \log(mn) \log(2m/n)})$.

Another possible approach is to bound the length of vector A_i^S when each entry in the vector is from $N(0, \sigma^2)$ (without bounding the maximum coefficient): using Lemma 4, for any fixed $S \subseteq [n]$ and $i \in [m]$,

$$\Pr\left(\|A_i^S\|^2 - |S|\sigma^2 \geq \lambda\sigma^2\right) \leq 2e^{-\frac{\lambda^2|S|}{24}}.$$

By union bound, we get that

$$\begin{aligned} \Pr\left(\exists S \subseteq [n], i \in [m] : \|\|A_i^S\|^2 - |S|\sigma^2\| \geq \lambda\sigma^2\right) &\leq 2e^{-\frac{\lambda^2}{24|S|}} \cdot \binom{n}{|S|} \cdot m \\ &\leq 2e^{-\frac{\lambda^2}{24|S|}} n^{|S|} m. \end{aligned}$$

Thus, taking $\lambda = |S|\sqrt{48(\log(n) + (1/|S|)\log m)}$ we get $\|A_i^S\| \leq 48\sigma|S|\sqrt{\log n + (1/|S|)\log m}$ for every $i \in [m]$ and $S \subseteq [n]$ whp. Therefore, the discrepancy incurred in each round of fixing is at most

$$196\sigma|S|\sqrt{\left(\log n + \frac{1}{|S|}\log m\right) \log \frac{2m}{|S|}}$$

and the total bound on discrepancy of A^S is at most

$$O\left(\sigma|S|\sqrt{\log \frac{2m}{|S|}} \left(\sqrt{\log n} + \sqrt{\frac{\log m}{|S|}}\right)\right).$$

This bound is still large (e.g., this gives the total discrepancy to be at most $O(\sigma n \sqrt{\log n \log(2m/n)})$). In fact, when each entry is from $N(0, \sigma^2)$, it is possible that there exists a subset of coordinates $S \subseteq [n]$ such that the length of a_i^S is $\Omega(\sigma|S|)$.

However, in order to bound the total discrepancy, we only need to bound the length of the remaining vector after each round of fixing. Let S denote the set of coordinates to be fixed in the current round. The existence lemma (Lemma 6) picks some subset from S of at least $|S|/2$ coordinates to fix so that the discrepancy is at most $4\|A_i^S\| \sqrt{\log(2m/|S|)}$. Hence, it leaves at most $|S|/2$ coordinates among the possible $|S|$ coordinates for the next round. It is sufficient to bound the probability that there exists a subset $T \subseteq S$ of size at most $|S|/2$ such that the length of the vector A_i^T is large. We do not need the length of A_i^T to be small for every subset $T \subseteq [n]$. Thus, the union bound is only over the choices of the coordinates yet to be fixed (subsets of S of size at most $|S|/2$) and not over all possible subsets of coordinates. We use this approach in Lemma 10 to obtain a stronger bound on the length of the vectors $A_i^{S_k}$ for every $i \in [m]$ and every collection of subsets (S_1, S_2, \dots, S_k) where $S_k \subseteq S_{k-1}$ and $|S_k| \leq n2^{-k}$. This helps us obtain the tighter bound for hereditary discrepancy as stated in Theorem 6.

3.1 Bucket entropy for matrices

We follow the outline of the entropy method by Spencer [20] to derive the following bound.

Lemma 6. *For any set of vectors $A_1, \dots, A_m \in \mathbb{R}^n$ and any subset $S \subseteq [n]$, there exists a point $z \in \{0, -1, +1\}^{|S|}$ with at least $|S|/2$ non-zero coordinates such that*

$$|A_i^S z| \leq 8 \|A_i^S\| \sqrt{\log \frac{2m}{|S|}} \quad \forall i \in [m].$$

Our proof of Lemma 6 is nearly identical to the known classical proof. However, since we use a different normalization ($\|a\|$ rather than $\|a\|_\infty$), the statement we need is not a direct corollary and we have to formally run through the proof for completeness.

3.1.1 Proof idea

The proof is by the probabilistic method. We show that there exist two vectors $x, y \in \{+1, -1\}^{|S|}$ such that

1. $|A_i^S x - A_i^S y|$ is small for every $i \in [m]$
2. x and y differ in a large number of coordinates.

Thus, taking $z = \frac{x-y}{2}$ gives a vector z so that $z \in \{0, -1, +1\}^{|S|}$ and z has a large number of non-zero coordinates. Further, since $|A_i^S(x - y)|$ is small, $|A_i^S z|$ is also small for every $i \in [m]$.

In order to show that there exist vectors $x, y \in \{-1, +1\}^{|S|}$ satisfying condition 1 above, we consider the value $|A_i^S x|$ for every $x \in \{-1, +1\}^{|S|}$. We show that there exist $x, y \in \{-1, +1\}^{|S|}$ so that the difference between $|A_i^S x|$ and $|A_i^S y|$ is small for each $i \in [m]$. For this, we consider a real line for each $i \in [m]$ and equi-partition the i 'th line into small parts for each $i \in [m]$. Then, we show that there exist an exponential number of vectors $x \in \{-1, +1\}^{|S|}$ such that their corresponding $|A_i^S x|$ values fall in the same part for every $i \in [m]$. Thus, we get a set containing exponential number of vectors in $\{-1, +1\}^{|S|}$ so that for any pair of vectors x, y in this set, $|A_i^S x| - |A_i^S y|$ is at most the length of each part corresponding to $i \in [m]$. Therefore, we have an exponential number of vectors satisfying condition 1.

Finally, since an exponential number of vectors $x \in \{-1, +1\}^{|S|}$ satisfy condition 1, there should exist at least two such vectors x and y with large hamming distance. Thus, among the set of vectors satisfying property 1, there should exist at least two vectors satisfying property 2.

Notation. We define the following function for equi-partitioning. For any $\lambda > 0$, define buckets

$$\begin{aligned} B_0^\lambda &:= [-\lambda, \lambda] \\ \text{for every positive integer } l, B_l^\lambda &:= ((2l-1)\lambda, (2l+1)\lambda] \\ B_{-l}^\lambda &:= [-(2l+1)\lambda, -(2l-1)\lambda]. \end{aligned}$$

Suppose we have a real vector $t = (t_1, \dots, t_m)$. Then define the bucketing function $P^t(x) = (P_1^{t_1}(x), \dots, P_m^{t_m}(x))$ where

$$P_i^\lambda(x) = j \quad \text{if} \quad \sum_{j \in S} A_{ij} x_j \in B_j^\lambda.$$

Thus, the length of each part in the equipartition for the i 'th vector is $2t_i$.

3.1.2 Proof of Lemma 6

The following lemma shows the existence of a point z with discrepancy of A_i due to z being at most t_i for each $i \in [m]$ if the entropy of the bucketing function is small.

Lemma 7. *Suppose $x \in \{-1, +1\}^n$ is chosen uniformly at random. If $\text{ENT}(P^t(x)) \leq |S|/5$, then there exists a point $z \in \{0, -1, +1\}^{|S|}$ with at least $|S|/2$ non-zero coordinates such that*

$$|A_i^S z| \leq t_i \text{ for each } i \in [m].$$

Proof of Lemma 7. Let $r = |S|$. Since $\text{ENT}(P^t(x)) \leq \frac{r}{5}$, there exists a vector $b = (b_1, \dots, b_m)$ such that $\Pr(P^t(x) = b) \geq 2^{-\frac{r}{5}}$. Since total number of possible choices for x is 2^r , at least $2^{\frac{4r}{5}}$ of the choices for x should map to b . This implies that there exist x, y which differ in at least $r/2$ coordinates such that $P^t(x) = P^t(y)$ [13]. Taking $z = \frac{x-y}{2}$ completes the proof of the lemma. \square

We use the following lemma to bound the entropy. This is very similar to Lemma 2.3 in [16].

Lemma 8. *Let S be an arbitrary subset of $[n]$. Let $x \in \{-1, +1\}^{|S|}$ be chosen uniformly at random. Then $\text{ENT}(P_i^{t_i}(x)) \leq G(t_i / \|A_i^S\|)$ for every $i \in [m]$, where*

$$G(\lambda) = \begin{cases} 40e^{-\frac{\lambda^2}{9}} & \text{if } \lambda > 0.1, \\ 40 \ln(1/\lambda) & \text{if } \lambda \leq 0.1. \end{cases}$$

Proof of Lemma 8. Suppose we pick x uniformly at random in $\{+1, -1\}^{|S|}$. Let

$$p_k := \Pr(P_i^{\lambda \|A_i^S\|}(x) = k).$$

Then, $\text{ENT}(P_i^{\lambda \|A_i^S\|}(x)) = \sum_k -p_k \log(p_k)$. Also,

$$\begin{aligned} \mathbb{E}(x_j) &= 0 \text{ for each } j \in [n], \\ \mathbb{E}(x_j^2) &= 1 \text{ for each } j \in [n], \\ \mathbb{E}(A_i^S x) &= 0 \text{ for any } A_i, \\ \mathbb{E}((A_i^S x)^2) &= \sum_{j \in S} A_{ij}^2 = \|A_i^S\|^2 \text{ for any } A_i. \end{aligned}$$

By Lemma 5,

$$\Pr(A_i^S x \geq \lambda \|A_i^S\|) \leq e^{-\frac{\lambda^2}{2}}.$$

Define

$$\begin{aligned} g_k &:= e^{-\frac{\lambda^2(2k-1)^2}{8}}, \quad k \geq 1 \\ g_0 &:= 1 - 2e^{-\frac{\lambda^2}{8}}. \end{aligned}$$

By Lemma 5, $p_k, p_{-k} \leq g_k$ and $p_0 \geq g_0$. The function $-x \log x$ is increasing in $(0, 1/e)$ and decreasing in $[1/e, 1]$.

When $\lambda \geq 10$, $g_0 \geq 1/e$ and $g_k < 1/e$ for $k \geq 1$. Therefore,

$$\begin{aligned} \text{ENT} \left(P_i^{\lambda \|A_i^S\|}(x) \right) &\leq -g_0 \log g_0 + 2 \sum_{k=1}^{\infty} -g_k \log g_k \\ &\leq 26e^{-\frac{\lambda^2}{9}}. \end{aligned}$$

When $0.1 \leq \lambda \leq 10$, by Jensen's inequality, $\sum_{k=-100}^{100} -p_k \log p_k \leq \log |K| \leq 8$. For $|k| \geq 101$, $g_k < 1/e$ and hence

$$\sum_{k=101}^{\infty} -p_k \log p_k \leq \sum_{k=101}^{\infty} -g_k \log g_k \leq \frac{1}{2}.$$

Thus,

$$\text{ENT} \left(P_i^{\lambda \|A_i^S\|}(x) \right) \leq 9 \leq 26e^{-\frac{\lambda^2}{9}}.$$

When $\lambda < 0.1$, by Jensen's inequality, $\sum_{k: |k| \leq \lambda^{-20}} -p_k \log p_k \leq \log |K|$. For $|k| > \lambda^{-20}$, $g_k < 1/e$. Therefore,

$$\begin{aligned} \text{ENT} \left(P_i^{\lambda \|A_i^S\|}(x) \right) &\leq \log(1 + 2\lambda^{-20}) + 2 \sum_{k: |k| \geq \lambda^{-20}} -g_k \log g_k \\ &\leq 40 \ln(1/\lambda). \end{aligned}$$

□

Proof of Lemma 6. Let $r = |S|$. Suppose we pick x uniformly at random in $\{+1, -1\}^{|S|}$. We show that $\text{ENT}(P^t(x)) \leq r/5$ for $t_i = 8 \|A_i^S\| \sqrt{\log \frac{2m}{r}}$, $i \in [m]$. The existence of a point $z \in \{0, -1, +1\}^{|S|}$ with at least $|S|/2$ non-zero coordinates such that

$$|A_i^S z| \leq t_i \text{ for each } i \in [m]$$

follows by Lemma 7.

By sub-additivity of entropy function,

$$\text{ENT}(P^t(x)) \leq \sum_{i=1}^m \text{ENT}(P_i^{t_i}(x)).$$

Due to the choice of t_i , we have that $t_i / \|A_i^S\| = 8 \sqrt{\log(2m/r)} > 0.1$. Therefore, by Lemma 8,

$$\text{ENT}(P_i^{t_i}(x)) \leq 40e^{-(64/9) \log \frac{2m}{r}}.$$

Thus,

$$\text{ENT}(P^t(x)) \leq 40me^{-7 \log \frac{2m}{r}} \leq \frac{r}{5}.$$

□

3.2 Bounding lengths of Gaussian subvectors

Lemma 9. *If each entry A_{ij} is drawn i.i.d. from $N(0, \sigma^2)$, then with high probability*

$$\max_{i \in [m]} \|A_i^S\| \leq 2\sigma \sqrt{|S| \log mn}$$

for every subset $S \subseteq [n]$.

Proof. If each entry A_{ij} is drawn i.i.d. from $N(0, \sigma^2)$, then by Lemma 2 the maximum entry $|A_{ij}|$, $i \in [m], j \in [n]$ is at most $2\sigma\sqrt{\log mn}$ with high probability. \square

Next we obtain a bound on the length of A_i^S when $|S|$ is large.

Lemma 10. *Suppose we have a matrix $A \in \mathbb{R}^{m \times n}$ where $n \geq \log m$ and each entry A_{ij} is drawn from $N(0, \sigma^2)$. For any collection of subsets $S_0, S_1, S_2, \dots, S_{\log(n/\log m)}$ of the set $[n]$, where $S_0 \subseteq [n]$, $S_k \subseteq S_{k-1}$, $|S_k| \leq n2^{-k}$ for $k = 0, 1, \dots, \log(n/\log m)$, the following holds with high probability.*

$$\|A_i^{S_k}\|^2 \leq 16n2^{-k}\sigma^2 \tag{1}$$

for every $i \in [m]$ and $k = 0, 1, \dots, \log(n/\log m)$.

Proof of Lemma 10. Let S be a collection of subsets $S_0, S_1, \dots, S_{\log(n/\log m)}$ of $[n]$ such that $S_0 \subseteq [n]$, $S_k \subseteq S_{k-1}$ and $|S_k| \leq n2^{-k}$ for $k = 0, 1, \dots, \log(n/\log m)$. We will show that (1) holds for every possible S .

We say that a subset S_k of the collection S is *heavy* if there exists $i \in [m]$ such that it violates (1). We denote a collection S of subsets to be *heavy* if there exists $k \in \{0, 1, \dots, \log(n/\log m)\}$ such that S_k is heavy.

Thus, a collection is heavy if one of its subsets is heavy. We will bound the probability that there exists a heavy collection. Therefore,

$$\Pr(\exists S : S \text{ is heavy}) \leq \sum_{k=0}^{\log(n/\log m)} \Pr\left(\exists S_k \subseteq S_{k-1}, |S_k| \leq n2^{-k} : S_k \text{ is heavy}\right)$$

We bound each term in the above sum as follows. For $k = 0$,

$$\begin{aligned} \Pr(S_0 \subseteq [n] \text{ is heavy}) &\leq \Pr\left(\exists S_0 \subseteq [n], i \in [m] : \|A_i^{S_0}\|^2 > 16n\sigma^2\right) \\ &\leq \Pr\left(\exists i \in [m] : \|A_i\|^2 > 16n\sigma^2\right) \\ &\leq 2e^{-5n} \cdot m \quad (\text{Using Lemma 4}) \\ &\leq 2e^{-4 \log m} \quad (n \geq \log m). \end{aligned}$$

For each $k = 1, 2, \dots, \log(n/\log m)$,

$$\Pr\left(\exists S_k \subseteq S_{k-1}, |S_k| \leq n2^{-k} : S_k \text{ is heavy}\right)$$

$$\begin{aligned}
&\leq \Pr\left(\exists S_k \subseteq S_{k-1}, |S_k| \leq n2^{-k}, i \in [m] : \|A_i^{S_k}\|^2 > 16n2^{-k}\sigma^2\right) \\
&\leq 2e^{-\frac{256n^2 2^{-2k}}{48|S_k|}} \cdot 2^{n2^{-(k-1)}} \cdot m \quad (\text{Using Lemma 4}) \\
&\leq 2e^{-\left(\frac{5n}{2^k} - \frac{n}{2^{k-1}} - \log m\right)} \quad (|S_k| \leq n2^{-k}) \\
&\leq 2e^{-2\log m} \quad (k \leq \log(n/\log m)).
\end{aligned}$$

Thus,

$$\Pr(\exists S : S \text{ is heavy}) \leq \frac{2}{m^2} \cdot \log\left(\frac{n}{\log m}\right) \rightarrow 0.$$

□

3.3 Proof of Theorem 2

We prove Theorem 6 showing a bound on the hereditary discrepancy.

Proof of Theorem 6. We use Lemma 6 repeatedly to fix the coordinates of x . We start with $S_0 = S$. By Lemma 6 there exists a point $z_0 \in \{0, -1, +1\}^{|S_0|}$ containing at most $|S_0|/2$ zeros. Let S_1 denote the subset of coordinates of z_0 that are zero. Then we set $z(j) = z_0(j)$ for every $j \notin S_1$. We take $S = S_1$. By Lemma 6 there exists a point $z_1 \in \{0, -1, +1\}^{|S|}$ containing at most $|S|/2$ zeros. Let S_2 denote the subset of coordinates of z_1 that are zero. Then we set $x(j) = z_1(j)$ for every $j \notin S_2$. We repeat this until the number of coordinates of x that are yet to be set is at most a constant with high probability. We set these remaining coordinates to be $-1/1$ arbitrarily. The discrepancy incurred by x due to this arbitrary setting is at most a constant.

We use Lemma 10 to bound the discrepancy incurred when the number of coordinates to be fixed is greater than $\log m$ and Lemma 9 to bound the discrepancy incurred when the number of coordinates to be fixed is at most $\log m$.

By Lemma 6, the discrepancy incurred by x while setting its coordinates using subset S_k , $k \in \{0, 1, \dots, \log(n/\log m)\}$ is at most

$$\begin{aligned}
|A_i^{S_k} z_k| &\leq 8 \|A_i^{S_k}\| \sqrt{\log \frac{2m}{|S_k|}} \\
&\leq 32\sigma \sqrt{n2^{-k} \log \frac{2m}{n2^{-k}}}
\end{aligned}$$

with high probability. Here, the second inequality is by using Lemma 10.

Thus, the discrepancy incurred by x due to $z_0, z_1, \dots, z_{\log(n/\log m)}$ is at most

$$\begin{aligned}
\sum_{k=0}^{\log \frac{n}{\log m}} |A_i^{S_k} z_k| &\leq \sum_{k=0}^{\log \frac{n}{\log m}} 32\sigma \sqrt{n2^{-k} \log \frac{2m}{n2^{-k}}} \\
&\leq 32\sigma \sqrt{2n \log \frac{2m}{n}}
\end{aligned}$$

with high probability.

For $k \geq \log(n/\log m)$, the number of coordinates $|S_k| \leq \log m$. By Lemma 6, the discrepancy incurred by x while setting its coordinates using subset S_k , $k \in \{\log(n/\log m) + 1, \dots, \log |S|\}$ is at most

$$\begin{aligned} |A_i^{S_k} z_k| &\leq 8 \|A_i^{S_k}\| \sqrt{\log \frac{2m}{|S_k|}} \\ &\leq 16\sigma \sqrt{n2^{-k} \log(mn) \log \frac{2m}{n2^{-k}}} \end{aligned}$$

with high probability. Here, the second inequality is by using Lemma 9 and $|S_k| \leq n2^{-k}$.

Thus, the discrepancy incurred by x due to $z_{\log(n/\log m)+1}, \dots, z_{\log |S|}$ is at most

$$\begin{aligned} \sum_{k=\log \frac{n}{\log m}}^{k=\log |S|} |A_i^{S_k} z_k| &\leq \sum_{k=\log \frac{n}{\log m}}^{k=\log |S|} 16\sigma \sqrt{n2^{-k} \log(mn) \log \frac{2m}{n2^{-k}}} \\ &\leq 32\sigma \sqrt{\log m \log(mn) \log \frac{2m}{\log m}} \end{aligned}$$

with high probability. Hence, the total discrepancy is bounded by

$$\max_{i \in [n]} |A_i^S x| \leq 64\sigma \left(\sqrt{n \log \frac{2m}{n}} + \sqrt{\log m \log(mn) \log \frac{2m}{\log m}} \right)$$

with high probability. □

Finally, we bound the linear discrepancy of random matrix (Theorem 2). Theorem 2 follows from Theorems 5 and 6. We give a direct proof here for the sake of completeness. Our proof strategy is well-known (see Corollary 8 in [20]).

Proof of Theorem 2. We will find x by rounding x_0 . Without loss of generality, let x_0 be such that $x_0(j) \in [0, 1]$ for each $j \in [n]$. Let the vector x_0 be rational. Suppose each coordinate in x_0 can be expressed in binary using at most p bits. We will round in p phases - each phase will reduce the number of bits needed to express each coordinate in the rounded vector by one.

Consider the binary expansion $x_0(j) = \sum_{k=0}^p \delta_{j,k} 2^{-k}$, $\delta_{j,k} \in \{0, 1\}$. Let S denote the set of coordinates of x_0 which require precision at the p -th bit, i.e., $S = \{j : \delta_{j,p} = 1\}$. Now, by Theorem 6, there exists a point $z \in \{-1, +1\}^{|S|}$ such that

$$|A_i^S z| \leq 64\sigma \left(\sqrt{|S| \log \frac{2m}{|S|}} + \sqrt{\log m \log(m|S|) \log \frac{2m}{\log m}} \right).$$

Now, consider the following rounding procedure to obtain x_1 : Set $z(j) = 0$ for every $j \notin S$ and $x_1 = x_0 + z2^{-p}$. It is clear that the number of bits needed to express x_1 is at most $p - 1$. This is because, exactly those coordinates which required precision at the p -th bit were rounded. Further, they were rounded in a manner so that the p -th bit is set to 0. This is because $z(j) \in \{+1, -1\}$

for every $j \in S$ (rounding could possibly change the $p-1$ -th bit in each coordinate). We also have that

$$\begin{aligned} |A_i(x_1 - x_0)| &= |A_i z 2^{-p}| \\ &\leq 64\sigma \left(\sqrt{|S| \log \frac{2m}{|S|}} + \sqrt{\log m \log(m|S|) \log \frac{2m}{\log m}} \right) \cdot 2^{-p}. \end{aligned}$$

We repeat this rounding procedure at most $p-1$ times thereby reducing the number of bits of precision by at least one each time. Thus, the final \bar{x} obtained needs one bit of precision for each coordinate and hence $\bar{x} \in \{0, 1\}^n$. Finally,

$$\begin{aligned} |A_i(\bar{x} - x_0)| &\leq \sum_{k=1}^{p-1} |A_i(x_k - x_{k-1})| \\ &\leq 64\sigma \sum_{k=1}^p 2^{-k} \left(\sqrt{|S_k| \log \frac{2m}{|S_k|}} + \sqrt{\log m \log(mn) \log \frac{2m}{\log m}} \right) \\ &\leq 64\sigma \sum_{j=1}^p 2^{-j} \left(\sqrt{n \log \frac{2m}{n}} + \sqrt{\log m \log(mn) \log \frac{2m}{\log m}} \right) \quad (|S_k| \leq n) \\ &\leq 64\sigma \left(\sqrt{n \log \frac{2m}{n}} + \sqrt{\log m \log(mn) \log \frac{2m}{\log m}} \right). \end{aligned}$$

□

4 Proof of the main existence theorem

The upper bound R_1 for the radius in Theorem 1 will follow from the linear discrepancy bound given in Theorem 2. For the lower bound, we use the following result for Gaussian matrices.

Lemma 11. *For $m \geq 1000n$, let $A \in \mathbb{R}^{m \times n}$ be a matrix whose entries are chosen i.i.d. from the normal distribution $N(0, \sigma^2)$. Let $x_0 := (1/2, \dots, 1/2) \in \mathbb{R}^n$. Then,*

$$\Pr \left(\exists x \in \mathbb{Z}^n : |A_i(x - x_0)| \leq \frac{\sigma}{2} \sqrt{n \log \frac{2m}{n}} \quad \forall i \in [m] \right) \leq \frac{1}{2^n}.$$

We first show a bound on the radius required so that the random IP $P(n, m, 0, R)$ contains an integer point with all nonzero coordinates. Lemma 11 follows from the choice of x_0 .

Lemma 12. *For $m \geq 1000n$, let $A \in \mathbb{R}^{m \times n}$ be a matrix whose entries are chosen i.i.d. from the normal distribution $N(0, \sigma^2)$. With probability at least $1 - 2^{-n}$, there does not exist $x \in \mathbb{Z}^n \cap \{x \in \mathbb{R}^n : |x_j| \geq 1 \quad \forall j \in [n]\}$ such that*

$$|A_i x| \leq \sigma \sqrt{n \log(2m/n)} \quad \text{for every } i = 1, \dots, m.$$

Proof of Lemma 12. For each $r > 0$, we define the set

$$U_r := \mathbb{Z}^n \cap \{x : \|x\| = r, |x_j| > 0 \ \forall j \in [n]\}.$$

We will show that with probability at least $1 - 2^{-n}$, there does not exist $x \in \cup_{r \geq 0} U_r$ satisfying all the m inequalities. We first observe that U_r is non-empty only if $r \geq \sqrt{n}$. Fix $r \geq \sqrt{n}$ and a point $X \in U_r$. Now, for $i \in [m]$, since each A_{ij} is chosen from $N(0, \sigma^2)$, the dot product

$$A_i x = \sum_{j=1}^n A_{ij} x_j$$

is distributed according to the normal distribution $N(0, r^2 \sigma^2)$. Let

$$P_x := \Pr \left(|A_i x| \leq \sigma \sqrt{n \log \frac{2m}{n}} \ \forall i \in [m] \right),$$

$$P_r := \Pr \left(\exists x \in U_r : |A_i x| \leq \sigma \sqrt{n \log \frac{2m}{n}} \ \forall i \in [m] \right).$$

By union bound,

$$P_r \leq \sum_{x \in U_r} P_x \leq |U_r| \max_{x \in U_r} P_x.$$

We will obtain an upper bound on P_x that depends only on r . To bound the size of the set U_r , we observe that every point in U_r is an integer point on the surface of a sphere of radius r centered around the origin and hence is contained in an euclidean ball of radius $r + 1$ centered around the origin. Thus, $|U_r|$ can be bounded by the volume of the sphere of radius $r + 1 \leq 2r$ centered around the origin:

$$|U_r| \leq \text{vol}(2r \mathbb{B}_0) \leq \left(2r \sqrt{\frac{2\pi e}{n}} \right)^n \leq \left(\frac{10r}{\sqrt{n}} \right)^n.$$

Next we bound P_r . We have two cases.

Case 1. Let $r \in [\sqrt{n}, \sqrt{n \log(2m/n)}]$. Since $A_i x$ is distributed according to $N(0, r^2 \sigma^2)$, by Lemma 1,

$$\Pr \left(|A_i x| \leq \sigma \sqrt{n \log \frac{2m}{n}} \right) \leq 1 - \frac{1}{\sqrt{2\pi}} \left(\frac{r \sqrt{n \log \frac{2m}{n}}}{r^2 + n \log \frac{2m}{n}} \right) \cdot \left(\frac{n}{2m} \right)^{\frac{n}{2r^2}}.$$

Since each A_{ij} is chosen independently, we have that

$$\begin{aligned} P_x &= \prod_{i=1}^m \Pr \left(|A_i x| \leq \sigma \sqrt{n \log \frac{2m}{n}} \right) \\ &< \left(1 - \frac{1}{\sqrt{2\pi}} \left(\frac{r \sqrt{n \log \frac{2m}{n}}}{r^2 + n \log \frac{2m}{n}} \right) \cdot \left(\frac{n}{2m} \right)^{\frac{n}{2r^2}} \right)^m \\ &\leq e^{-\frac{1}{\sqrt{2\pi}} \left(\frac{r \sqrt{n \log \frac{2m}{n}}}{r^2 + n \log \frac{2m}{n}} \right) \cdot \left(\frac{n}{2m} \right)^{\frac{n}{2r^2}} \cdot m}. \end{aligned}$$

Therefore, by union bound, it follows that

$$\begin{aligned} P_r &\leq e^{-\frac{1}{\sqrt{2\pi}} \left(\frac{r\sqrt{n \log \frac{2m}{n}}}{r^2 + n \log \frac{2m}{n}} \right) \cdot \left(\frac{n}{2m} \right)^{\frac{n}{2r^2}} \cdot m + n \log \frac{10r}{\sqrt{n}}} \\ &\leq e^{-n \log \frac{10r}{\sqrt{n}}} \leq \left(\frac{\sqrt{n}}{10r} \right)^n. \end{aligned}$$

Case 2. Let $r > \sqrt{n \log (2m/n)}$. Since $A_i x$ is distributed according to $N(0, r^2 \sigma^2)$, by Lemma 1, we have that

$$\Pr \left(|A_i x| \leq \sigma \sqrt{n \log \frac{2m}{n}} \right) \leq \frac{1}{r} \sqrt{\frac{2}{\pi} n \log \frac{2m}{n}} \leq \frac{4}{5r} \sqrt{n \log \frac{2m}{n}}.$$

The random variables $A_1 x, \dots, A_m x$ are independent and identically distributed. Therefore,

$$\begin{aligned} P_x &= \prod_{i=1}^m \Pr \left(|A_i x| \leq \sigma \sqrt{n \log \frac{2m}{n}} \right) \\ &\leq \left(\frac{4}{5r} \sqrt{n \log \frac{2m}{n}} \right)^m. \end{aligned}$$

Hence, by union bound,

$$\begin{aligned} P_r &\leq e^{-n \left(\frac{m}{n} \log \left(\frac{5r}{4\sqrt{n \log \frac{2m}{n}}} \right) - \log \frac{10r}{\sqrt{n}} \right)} \\ &\leq e^{-n \left(\frac{m}{2n} \log \left(\frac{5r}{4\sqrt{n \log \frac{2m}{n}}} \right) \right)} \\ &\leq \left(\frac{4\sqrt{n \log \frac{2m}{n}}}{5r} \right)^{\frac{m}{2}}. \end{aligned}$$

Finally,

$$\begin{aligned} \Pr \left(\exists x \in \cup_{r \geq \sqrt{n}} U_r : |A_i x| \leq \sigma \sqrt{n \log \frac{2m}{n}} \ \forall i \in [m] \right) &= \sum_{r \geq \sqrt{n}} P_r \\ \sum_{r \geq \sqrt{n}} P_r &= \sum_{r \in [\sqrt{n}, \sqrt{n \log \frac{2m}{n}}]} P_r + \sum_{r > \sqrt{n \log \frac{2m}{n}}} P_r \\ &\leq \frac{1}{10^n} \int_{r=\sqrt{n}}^{\infty} \left(\frac{\sqrt{n}}{r} \right)^n dr + \left(\frac{4}{5} \right)^{\frac{m}{2}} \int_{r=\sqrt{n \log \frac{2m}{n}}}^{\infty} \left(\frac{\sqrt{n \log \frac{2m}{n}}}{r} \right)^{\frac{m}{2}} dr \\ &\leq \frac{1}{10^n} \cdot \frac{\sqrt{n}}{n-1} + \left(\frac{4}{5} \right)^{\frac{m}{2}} \cdot \left(\frac{2\sqrt{n \log \frac{2m}{n}}}{m-2} \right) \\ &\leq \frac{1}{2^n} \quad (\text{since } m \geq 1000n). \end{aligned}$$

□

Proof of Lemma 11. There exists $x \in \mathbb{Z}^n$ such that

$$|A_i(x - x_0)| \leq \frac{\sigma}{2} \sqrt{n \log \frac{2m}{n}} \quad \forall i \in [m]$$

if and only if there exists $x \in \mathbb{Z}^n \cap \{x \in \mathbb{R}^n : x_j \geq 1 \quad \forall j \in [n]\}$ such that

$$|A_i x| \leq \sigma \sqrt{n \log \frac{2m}{n}} \quad \forall i \in [m].$$

The result follows by Lemma 12. □

We are now ready to put everything together and prove Theorem 1.

Proof of Theorem 1. Let

$$P = \{x \in \mathbb{R}^n : a_i x \leq b_i \quad \forall i \in [m]\}$$

where each a_i is chosen from a spherically symmetric distribution. Then $\alpha_i = a_i / \|a_i\|$ for $i \in [m]$ is distributed randomly on the unit sphere. A random unit vector α_i can be obtained by drawing each coordinate from the normal distribution $N(0, \sigma^2 = 1/n)$ and normalizing the resulting vector. Thus, we may assume $\alpha_i = A_i / \|A_i\|$ where each coordinate A_{ij} is drawn from the normal distribution $N(0, 1/n)$. Here, we show that the probability that there exists a vector A_i that gets scaled by more than a constant is at most $2me^{-n/96}$.

Taking $r = n$ and $\sigma^2 = 1/n$ in Lemma 4, we have

$$\Pr \left(\exists i \in [m] : |\|A_i\|^2 - 1| > \frac{1}{2} \right) \leq 2me^{-\frac{n}{96}}.$$

Hence, with probability at least $1 - 2me^{-n/96}$, we have that $\sqrt{1/2} \leq \|A_i\| \leq \sqrt{3/2}$ for every $i \in [m]$.

1. Since P contains a ball of radius R_1 , $P \supseteq Q$ where

$$Q = \{x \in \mathbb{R}^n : |\alpha_i(x - x_0)| \leq R_1 \text{ for } i \in [m]\}$$

Using Theorem 2 and $\sigma^2 = 1/n$, we know that there exists $x \in \mathbb{Z}^n$ such that for every $i \in [m]$

$$|A_i(x - x_0)| \leq 64 \left(\sqrt{\log \frac{2m}{n}} + \sqrt{\frac{\log m \log(mn)}{n} \log \frac{2m}{\log m}} \right).$$

Thus, with probability at least $1 - 2me^{-n/96}$, there exists $X \in \mathbb{Z}^n$ satisfying

$$\begin{aligned} |\alpha_i(x - x_0)| &= \frac{|A_i(x - x_0)|}{\|A_i\|} \\ &\leq 128 \left(\sqrt{\log \frac{2m}{n}} + \sqrt{\frac{\log m \log(mn)}{n} \log \frac{2m}{\log m}} \right) \end{aligned}$$

for every $i \in [m]$. Thus the polytope Q is integer feasible and consequently P is also integer feasible.
 2. For $x_0 = (1/2, \dots, 1/2)$, let

$$P = \left\{ X \in \mathbb{R}^n : |A_i(x - x_0)| \leq \|A_i\| \sqrt{\frac{1}{6} \log \frac{2m}{n}} \forall i \in [m] \right\}.$$

Then, P contains a ball of radius R_0 centered around x_0 and hence is an instance of the random polytope $P(n, m, x_0, R_0)$. Further, with probability at least $1 - 2me^{-n/96}$, P is contained in

$$Q = \left\{ x \in \mathbb{R}^n : |A_i(x - x_0)| \leq \frac{1}{2} \sqrt{\log \frac{2m}{n}} \forall i \in [m] \right\}.$$

By Lemma 11, with high probability, we have that $Q \cap \mathbb{Z}^n = \emptyset$. Thus, with probability at least $1 - 2me^{-n/96}$, we have that $P \cap \mathbb{Z}^n = \emptyset$. □

5 Algorithm to find an integer point

Next we present an algorithm to identify an integer point x in $P(n, m, x_0, R_{ALG})$. We first show an algorithm to find small linear discrepancy solutions for gaussian constraint matrices.

Theorem 7. *There is a randomized polynomial-time algorithm that takes as input a random matrix $A \in \mathbb{R}^{m \times n}$ with i.i.d. entries from $N(0, \sigma^2)$ and a point $x_0 \in \mathbb{R}^n$, and outputs an integer point x such that for every $i \in [m]$,*

$$|A_i(x - x_0)| \leq 2^{15} \sigma \left(\sqrt{n} \log \frac{2m}{n} + \sqrt{\log m \log(mn)} \log \frac{2m}{\log m} \right)$$

with high probability.

Our algorithm is similar to Bansal's algorithm [1]. The algorithm runs in phases. In each phase, we start with the current point $\bar{x} \in [0, 1]^n$ and perform a random walk to arrive at a partial vector y with at least half of the non-integer coordinates of \bar{x} being integers in y . Further, the discrepancy overhead incurred by y (i.e., $|A_i(y - \bar{x})|$) is small.

Algorithm Round-IP

Input: Point $x_0 \in \mathbb{R}^n$, matrix $A \in \mathbb{R}^{m \times n}$ where each $A_{ij} \sim N(0, \sigma^2)$.

Output: An integer point x in the polytope

$$P = \left\{ x : |A_i(x - x_0)| \leq 2^{15} \sigma \left(\sqrt{n} \log \frac{2m}{n} + \sqrt{\log m \log(mn)} \log \frac{2m}{\log m} \right) \right\}.$$

1. Initialize $x = x_0$.

2. While(x is not integral)

(a) Define S to be the set of non-integer coordinates of x .

(b) $x \leftarrow \text{Partial-Vector}(A, x, S)$.

(c) If $x = \text{NULL}$, then abort.

3. Output x .

We describe the partial vector function used in the above algorithm in the next section. Its functionality is summarized in the following lemma.

Lemma 13. *Given a point $x_0 \in [0, 1]^n$, let $S \subseteq [n]$ denote the subset of non-integer coordinates of x_0 . There exists a polynomial time algorithm that produces a vector $y \in [0, 1]^n$ with at most $|S|/2$ non-integer coordinates, such that*

$$|A_i(y - x_0)| \leq 1281 \|A_i^S\| \log \frac{2m}{|S|}$$

for every $i \in [m]$ with probability at least $1/2$.

In Algorithm Round-IP, we repeatedly invoke the Partial-Vector algorithm. Each such call fixes at least half the non-integer coordinates of x to integers. Thus, with at most $\log n$ calls to the partial-vector algorithm, we obtain an integer vector x . Further, the total discrepancy overhead incurred by x is at most the sum of the discrepancy overhead incurred in each call to the Partial-Vector algorithm. The sum of the discrepancy overheads is bounded similar to the proof of Theorem 6 using Lemmas 9 and 10.

5.1 Partial integer vector for Gaussian matrices

We show a polynomial time algorithm that given a point x_0 as input, finds a point y such that the number of non-integer coordinates is halved and the discrepancy overhead incurred by y is small. This algorithm is along the lines of Bansal's algorithm [1]. Bansal's algorithm outputs a partial vector with small discrepancy overhead when the matrix A is a 0/1 matrix. Here, we need an algorithm to find a partial vector with small discrepancy overhead when each entry in A is drawn i.i.d. from $N(0, \sigma^2)$.

The vector y is constructed iteratively by performing a random walk starting at x_0 . In the t 'th iteration, $x_t = x_{t-1} + \gamma_t$ for an appropriate choice of γ_t . Let S denote the set of non-integer coordinates of x_0 . The increments $\gamma_t(j)$ satisfy the following properties.

1. The increment $\gamma_t(j)$ is zero if $j \notin S$ and are distributed as an unbiased gaussian with standard deviation at most 1 if $j \in S$. Thus, each non-integer coordinate evolves as a martingale.
2. The increments are such that they add up to at least $s(|S|/2)$ for some scaling parameter s . This ensures that after about $1/s^2$ steps, about half the coordinates are integers (or close to integers).
3. Finally, at any step t , the increments $(\gamma_t(j))_{j \in S}$ are correlated such that $\sum_{j \in S} A_{ij} \gamma_t(j)$ is distributed as an unbiased Gaussian with small standard deviation (proportional to the length of A_i^S). This ensures that the discrepancy overhead incurred is also evolving as a martingale with small increments.

5.1.1 Algorithm

We describe the Partial-Vector algorithm with the following parameters: $r = |S|$, $u = \log m$, $s = 1/(4u^{3/2})$, $q = \log(2m/r)$ and for all $i \in [m]$,

$$\beta_i(0) := 0, \quad \beta_i(k) := 640q \|A_i^S\| \left(2 - \frac{1}{k}\right) \forall k = 1, 2, \dots, \quad \text{and} \quad \alpha_i(k) := \frac{100q \|A_i^S\|^2}{(k+1)^5} \forall k = 0, 1, 2, \dots$$

Algorithm Partial-Vector(A, x_0, S)

Input: Vector x_0 with index set S of non-integer coordinates and a matrix $A \in \mathbb{R}^{m \times n}$

Output: Vector x with at most $|S|/2$ non-integer coordinates such that

$$|A_i(x - x_0)| \leq 1281 \|A_i^S\| \log \frac{2m}{|S|} \text{ for every } i \in [m].$$

Initialize $C(0) \leftarrow S$.

Repeat for $t = 1, \dots, 16/s^2$:

1. Let

$$\gamma_\tau(A_i) := A_i(x_\tau - x_{\tau-1}), \quad \forall i \in [m], \quad \tau = 1, \dots, t-1$$

$$\eta_i := \left| \sum_{\tau=1}^{t-1} \gamma_\tau(A_i) \right| \quad \forall i \in [m], .$$

Declare a vector A_i to be k -dangerous if $\eta_i \in [\beta_i(k), \beta_i(k+1)]$. Let $S(k) := \{i : A_i \text{ is } k\text{-dangerous}\}$. If $\eta_i > 2\beta(1)$ for any $i \in [m]$, then abort.

2. Find a feasible solution to the following semidefinite program:

$$\begin{aligned} \sum_{j \in [n]} \|v_j\|^2 &\geq \frac{|C(t-1)|}{2} \\ \left\| \sum_{j=1}^n A_{ij} v_j \right\|^2 &\leq \alpha_i(k) \quad \forall i \in S(k), k = 0, 1, 2, \dots \\ \|v_j\|^2 &\leq 1 \quad \forall j \in C(t-1) \\ \|v_j\|^2 &= 0 \quad \forall j \notin C(t-1) \end{aligned}$$

If the SDP is infeasible, then abort.

3. Obtain each coordinate g_k according to the standard normal distribution $N(0, 1)$.

$$\begin{aligned} \gamma_t(j) &:= s \langle g, v_j \rangle, \\ x_t &\leftarrow x_{t-1} + \gamma_t. \end{aligned}$$

If $x_t(j) > 1$ or $x_t(j) < 0$ for any j , then abort.

4. Let $B(t) := \{j : x_t(j) \geq 1 - \frac{1}{u} \text{ or } x_t(j) < \frac{1}{u}\}$. For every $j \in B(t)$,

$$x_t(j) \leftarrow \begin{cases} 1 & \text{with probability } x_t(j), \\ 0 & \text{with probability } 1 - x_t(j). \end{cases}$$

5. Update $C(t) \leftarrow C(t-1) \setminus B(t)$. If $|C(t)| \leq |S|/2$, then terminate and output x_t .

5.1.2 Discrepancy incurred due to randomized rounding

We first show that the discrepancy overhead incurred due to randomized rounding in step 4 is small. Let x' denote the vector obtained if the rounding in Step 4 is not performed. That is, x' is the vector obtained at the end of the algorithm and whose respective coordinates were fixed once they exceeded $1 - (1/u)$ or become smaller than $1/u$.

Claim 14. *With high probability,*

$$|A_i(x - x')| \leq 4 \|A_i^S\| \quad \forall i \in [m].$$

Proof of Claim 14. First observe that $|A_i(x - x')| \leq |\sum_{j \in S} A_{ij}(x(j) - x'(j))|$. Consider a coordinate $j \in S$ which was rounded in step 4. Then,

$$\begin{aligned} \mathbb{E}(x(j) - x'(j)) &= 0 \\ \text{Var}(x(j) - x'(j)) &\leq \frac{1}{u}. \end{aligned}$$

Since the variables in S are the only variables that can get rounded,

$$\text{Var}\left(\sum_{j \in S} A_{ij}(x(j) - x'(j))\right) := \Delta_i^2 \leq \frac{\|A_i^S\|^2}{u}.$$

Therefore, for $i \in [m]$, by Chernoff bound,

$$\Pr\left(\left|\sum_{j \in S} A_{ij}(x(j) - x'(j))\right| \geq 4\Delta_i \sqrt{\log m}\right) = \frac{2}{m^2}.$$

Hence, by union bound, we get that $|A_i(x - x')| \leq 4\Delta_i \sqrt{\log m} \leq 4\|A_i^S\|$ for every $i \in [m]$ with high probability. \square

5.1.3 Feasibility of semidefinite program

Next we show that the SDP during iteration t is feasible with high probability.

Lemma 15. *Suppose the number of k -dangerous vectors is at most $r2^{-10(k+1)}$ for every $k = 1, 2, \dots$ during iteration t . Then the SDP in iteration t is feasible with high probability.*

Proof of Lemma 15. It is enough to prove that there exists a $z \in \{0, +1, -1\}^{|C(t-1)|}$ with at least $|C(t-1)|/2$ non-zero coordinates such that $|A_i^{C(t-1)} z| \leq \sqrt{\alpha_i(k)}$ exists for every k -dangerous vector A_i . Such a point z gives a feasible solution to the SDP. We show the existence of such a point z similar to the proof of Lemma 6.

By Lemma 7, it is sufficient to show that $\text{ENT}(P^\Lambda(x)) \leq r/5$, where $\lambda_i \leq \sqrt{\alpha_i(k)}$ if $i \in S(k)$. We take $\lambda_i = 10 \|A_i^{C(t-1)}\| \sqrt{(1/(k+1)^5) \log(2m/r)}$ for every $i \in S(k)$. By sub-additivity of entropy function

$$\text{ENT}(P^\Lambda(x)) \leq \sum_{i=1}^m \text{ENT}(P_i^{\lambda_i}(x)) = \sum_{k=0,1,\dots} \sum_{i \in S(k)} \text{ENT}(P_i^{\lambda_i}(x)).$$

Let $i \in S(k)$. Let $\zeta(k) := 8\sqrt{q/(k+1)^5}$. Since, $\zeta(0) = 10\sqrt{q} \geq 0.1$, by Lemma 8,

$$\sum_{i \in S(0)} \text{ENT} \left(P_i^{\lambda_i}(x) \right) \leq 40me^{-11q} \leq \frac{r}{10}.$$

Next, suppose $i \in S(k)$ such that $k \geq 1$. The function G in Lemma 8 is a decreasing function. Thus, if $\zeta(k) > 0.1$, then we can use $G(\zeta(k)) \leq 40 \ln(10)$ as an upper bound for $\text{ENT} \left(P_i^{\lambda_i}(x) \right)$. Therefore,

$$\text{ENT} \left(P_i^{\lambda_i}(x) \right) \leq 40 \max(\ln(10), \ln(1/\zeta(k))) \leq 40 \max(\ln(10), \ln((k+1)^{5/2})) \leq 200 \ln(k+1).$$

Thus,

$$\sum_{k \geq 1} \sum_{i \in S(k)} \text{ENT} \left(P_i^{\lambda_i}(X) \right) \leq \sum_{k \geq 1} r 2^{-10(k+1)} \cdot 200 \ln(k+1) \leq \frac{r}{10}.$$

Hence,

$$\text{ENT} (P^\Lambda(x)) \leq \frac{r}{5}$$

□

Next we bound the probability that a large number of vectors are k -dangerous. For this, observe that the increment $\gamma_t(j) = 0$ if $i \notin C(t-1)$ and $\gamma_t(j)$ is distributed according to the normal distribution $N(0, s^2)$ if $i \in C(t-1)$. Similarly, $\gamma_t(A_i)$ is distributed according to the normal distribution $N(0, \sigma'^2)$ where $\sigma'^2 \leq \alpha_i(k)s^2$ if $i \in S(k)$. This is due to the SDP constraint.

Lemma 16. *For $k = 1, 2, \dots$, let D_k denote the event that more than $m_k = r 2^{-10(k+1)}$ vectors ever become k -dangerous during $t = 1, 2, \dots, 16/s^2$. Then,*

$$\Pr(D_k) \leq 2^{-5(k+1)}.$$

Proof of Lemma 16. First consider $k = 1$. Suppose a vector A_i becomes 1-dangerous at some iteration t . Then, there exists \hat{t} when η_i first exceeds $\beta(1)$. Until \hat{t} , η_i was evolving as a martingale with each conditional increment drawn from Gaussian distribution with mean 0 and variance at most $\alpha(0)s^2$. Hence, by Lemma 3,

$$\Pr \left(\eta_i > \beta_i(1) \text{ at some } t = \hat{t} < \frac{16}{s^2} \right) \leq 2e^{-\frac{\beta_i(1)^2}{2\alpha_i(0)s^2(16/s^2)}} = 2 \left(\frac{r}{2m} \right)^{64}.$$

Hence, the expected number of vectors that become 1-dangerous is at most $m(r/2m)^{64} \leq r 2^{-30}$. By Markov's inequality

$$\Pr(|S(1)| \geq m_1) \leq \frac{r 2^{-30}}{r 2^{-20}} \leq 2^{-10}.$$

Now suppose $k \geq 2$. If A_i becomes k -dangerous during iteration t , then it was $(k-1)$ -dangerous at some iteration $\hat{t} < t$ and η_i increased by $\beta_i(k) - \beta_i(k-1)$ in at most $16/s^2$ iterations. Since $\text{Var}(\gamma_t(A_i)) \leq \alpha_i(k-1)s^2$ when $\eta_i \in [\beta_i(k-1), \beta_i(k)]$, we get that

$$\Pr \left(A_i \text{ becomes } k\text{-dangerous in at most } \frac{16}{s^2} \text{ iterations} \right) \leq 2e^{-\frac{(\beta_i(k) - \beta_i(k-1))^2}{4\alpha_i(k-1)s^2(16/s^2)}} \leq \frac{r}{2m} e^{-32k}.$$

Thus, expected number of k -dangerous vectors is at most $re^{-32k}/2 \leq r 2^{-15(k+1)}$. Hence, by Markov's inequality

$$\Pr(|S(k)| \geq m_k) \leq 2^{-5(k+1)}.$$

□

5.1.4 Proof of Lemma 13

We first show that the probability that the number of non-integer variables is at least $r/2$ after $16/s^2$ iterations is at most $1/4$. Then, we bound the probability that none of the events D_k happen. Conditioned on the event that none of the events D_k happen, we get that the SDP is feasible for every iteration t and the discrepancy incurred is bounded by at most $2\beta(1)$. Since the discrepancy overhead incurred by randomized rounding is small, we get Lemma 13.

The following Lemma is identical to that of Lemma 4.1 in [1]. We give a proof here for the sake of completeness.

Claim 17. *Let E be the event that the number of non-integer variables is at least $r/2$ after $16/s^2$ iterations. Then,*

$$\Pr(E) \leq \frac{1}{4}.$$

Proof of Claim 17. Define for every $t = 0, 1, 2, \dots, 16/s^2$,

$$w_t = \begin{cases} \sum_{j=1}^n x_t(j)^2 & \text{if } |C(t)| \geq \frac{r}{2} \\ w_{t-1} + \frac{s^2 r}{4} & \text{if } |C(t)| < \frac{r}{2}. \end{cases}$$

Now, if $|C(t)| < \frac{r}{2}$, then $w_t - w_{t-1} = (s^2 r/4)$. If $|C(t)| \geq r/2$, then

$$\begin{aligned} \mathbb{E}(w_t - w_{t-1}) &= \mathbb{E}\left(\sum_{j=1}^n ((x_{t-1}(j) + \gamma_t(j))^2 - \sum_{j=1}^n x_{t-1}(j)^2)\right) \\ &= \sum_{j=1}^n \mathbb{E}_g(\gamma_t(j)^2) \quad (\text{Since } \mathbb{E}_g(\gamma_t(j)) = 0) \\ &= \sum_{j \in C(t-1)} s^2 \\ &\geq \frac{s^2 r}{4}. \end{aligned}$$

Thus, $\mathbb{E}(w_t - w_{t-1}) \geq s^2 r/4$ for every $t = 1, 2, \dots, 16/s^2$. Further, if $|C(t)| \geq r/2$, then $w_t \leq r$, and hence, $w_t \leq r + (1/4)ts^2 r$ for every $t = 1, 2, \dots, 16/s^2$. Therefore, for $t_0 = 16/s^2$, we have that

$$\begin{aligned} \frac{t_0 s^2 r}{4} &\leq \mathbb{E}(w_{t_0} - w_0) \leq \mathbb{E}(w_{t_0}) \\ &\leq \Pr\left(|C(t_0)| \geq \frac{r}{2}\right) \cdot r + \Pr\left(|C(t_0)| < \frac{r}{2}\right) \cdot \left(r + \frac{t_0 s^2 r}{4}\right) \\ &= \Pr(E) r + (1 - \Pr(E)) \left(r + \frac{t_0 s^2 r}{4}\right) \\ &= \left(r + \frac{t_0 s^2 r}{4}\right) - \Pr(E) \frac{t_0 s^2 r}{4}. \end{aligned}$$

Thus, $\Pr(E) \leq 4/t_0 s^2 = 1/4$. □

Proof of Lemma 13. Since $\gamma_t(j)$ is distributed as Gaussian with variance at most s^2 , by Lemma 2, the algorithm does not abort in step (3) with high probability.

Let $D = \cup_{k=1}^{\infty} D_k$ and E be the event as defined in Claim 17. Now,

$$\Pr(D) \leq \sum_{k=1,2,3,\dots} \Pr(D_k) \leq \frac{1}{16}.$$

Thus, with probability at least $15/16$, \overline{D} holds. If \overline{D} holds, then the semidefinite program is feasible for every $t = 0, 1, \dots, 16/s^2$, and hence the algorithm does not abort. Further, if \overline{D} holds, then $m_k < 1$ for $k = 2 \log m$. Hence, there are no $(2 \log m)$ -dangerous vectors. Therefore, the discrepancy incurred for A_i is at most $\beta_i(2 \log m) + 4 \|A_i^S\|$. Therefore, the total discrepancy incurred is at most

$$\begin{aligned} |A_i(x_{t_0} - x_0)| &\leq \beta_i(2 \log m) + 4 \|A_i^S\| \\ &\leq 2\beta_i(1) + 4 \|A_i^S\| \\ &\leq 1281q \|A_i^S\|. \end{aligned}$$

Now, it is sufficient to show that $\Pr(\overline{D} \cap \overline{E}) \geq 1/2$. By Claim 17, we have that $\Pr(E) \leq 1/4$. Therefore,

$$\Pr(\overline{D} \cap \overline{E}) \geq 1 - \Pr(D) - \Pr(E) \geq \frac{1}{2}.$$

□

5.2 Finding an integer point

Proof of Theorem 7. Without loss of generality, we may assume that $x_0 \in [0, 1]^n$ and our objective is to find $x \in \{0, 1\}^n$ with low discrepancy overhead. We use Algorithm Round-IP. We will show that it succeeds with probability at least $1/n$ to find a point $x \in \{0, 1\}^n$ such that

$$|A_i(x - x_0)| \leq 2^{15} \sigma \left(\sqrt{n} \log \frac{2m}{n} + \sqrt{\log m \log mn} \log \frac{2m}{\log m} \right).$$

The success probability of this algorithm can be amplified by repeating it n times.

Let \overline{x} denote the vector output the Algorithm Round-IP and let x_k denote the vector x in Algorithm Round-IP after k calls to the Partial-Vector algorithm. Let S_k denote the set of non-integer coordinates in x_k . By Lemma 13, the discrepancy overhead incurred in the k 'th run of the random walk based Partial-Vector algorithm for $k \in \{0, 1, \dots, \log(n/\log m)\}$ is

$$\begin{aligned} |A_i^{S_k}(x_k - x_{k-1})| &\leq 1281 \|A_i^{S_k}\| \log \frac{2m}{|S_k|} \\ &\leq 2^{13} \sigma \sqrt{n 2^{-k}} \log \frac{2m}{n 2^{-k}} \end{aligned}$$

with high probability. Here, the second inequality is by using Lemma 10 and $|S_k| \leq n 2^{-k}$.

Thus, the discrepancy overhead incurred after $\log(n/\log m)$ calls to the Partial-Vector algorithm is

$$\begin{aligned} \sum_{k=0}^{\log \frac{n}{\log m}} |A_i^{S_k}(x_k - x_{k-1})| &\leq \sum_{k=0}^{\log \frac{n}{\log m}} 2^{14} \sigma \sqrt{n 2^{-k}} \log \frac{2m}{n 2^{-k}} \\ &\leq 2^{15} \sigma \sqrt{n} \log \frac{2m}{n} \end{aligned}$$

with high probability.

For $k \geq \log(n/\log m)$, the number of non-integer coordinates $|S_k| \leq \log m$. By Lemma 13, the discrepancy overhead incurred in the k 'th call to the Partial-Vector algorithm, where $k \in \{\log(n/\log m) + 1, \dots, \log n\}$ is

$$\begin{aligned} |A_i^{S_k}(x_k - x_{k-1})| &\leq 1281 \|A_i^{S_k}\| \log \frac{2m}{|S_k|} \\ &\leq 2^{12} \sigma \sqrt{n 2^{-k} \log(mn)} \log \frac{2m}{n 2^{-k}} \end{aligned}$$

with high probability. Here, the second inequality is by using Lemma 9 and $|S_k| \leq n 2^{-k}$.

Thus, the discrepancy overhead incurred by Algorithm Round-IP

$$\begin{aligned} \sum_{k=\log \frac{n}{\log m}}^{k=\log n} |A_i^{S_k}(x_k - x_{k-1})| &\leq \sum_{k=\log \frac{n}{\log m}}^{k=\log n} 2^{13} \sigma \sqrt{n 2^{-k} \log(mn)} \log \frac{2m}{n 2^{-k}} \\ &\leq 2^{14} \sigma \sqrt{\log m \log(mn)} \log \frac{2m}{\log m} \end{aligned}$$

with high probability.

Since each call to the Partial-Vector algorithm sets at least half of the remaining non-integer coordinates to integers, we call the Partial-Vector algorithm at most $\log n$ times. Each call succeeds with probability $1/2$. Hence, with probability at least $1/2^{\log n} = 1/n$, we obtain an integer point $\bar{x} \in \{0, 1\}^n$ such that the total discrepancy overhead is bounded as follows:

$$\max_{i \in [n]} |A_i(\bar{x} - x_0)| \leq 2^{15} \sigma \left(\sqrt{n} \log \frac{2m}{n} + \sqrt{\log m \log(mn)} \log \frac{2m}{\log m} \right).$$

□

Proof of Theorem 3. We use Theorem 7 to derive Theorem 3.

Let

$$R_{ALG} = 2^{16} \sigma \left(\sqrt{n} \log \frac{2m}{n} + \sqrt{\log m \log(mn)} \log \frac{2m}{\log m} \right)$$

and $\alpha_i = A_i / \|A_i\|$, $i \in [m]$. Solve the following linear programming problem to find the center of the largest ball contained in the polytope.

$$\begin{aligned} \max R \\ R \leq b_i - A_i x, \text{ for } i \in [m]. \end{aligned}$$

Let (x_0, R) be a solution to the above LP. Since P contains a ball of radius R_{ALG} , there exists $\beta_i \geq R_{ALG} \|A_i\|$ for every $i \in [m]$ such that

$$P \supseteq \{x \in \mathbb{R}^n \mid |A_i(x - x_0)| \leq \beta_i \text{ for } i \in [m]\}.$$

Observe that the polytope

$$Q = \{x \in \mathbb{R}^n \mid |\alpha_i(x - x_0)| \leq R_{ALG} \text{ for } i \in [m]\}$$

is contained in P . We will show that there exists a randomized polynomial-time algorithm to find an integer point in Q that succeeds with probability at least $(1 - 2me^{-n/96})/2$. This success probability is over the choice of A_i s.

Since each A_i is drawn from a spherically symmetric distribution, $\alpha_i = A_i / \|A_i\|$ is distributed uniformly on the unit sphere. A random unit vector α_i on a sphere is obtained by drawing each coordinate a_{ij} i.i.d. from the normal distribution $N(0, 1/n)$ and scaling the resulting vector by $\|a_i\| = \sqrt{\sum_{j=1}^n a_{ij}^2}$. Similar to the proof of Theorem 1, a_i gets scaled by at most 2 for every $i \in [m]$ with probability at least $1 - 2me^{-n/96}$. Using Theorem 7 and $\sigma^2 = 1/n$, we know that there exists a randomized polynomial time algorithm that succeeds with high probability to find $x \in \mathbb{Z}^n$ such that for every $i \in [m]$

$$|a_i(x - x_0)| \leq 2^{15} \left(\log \frac{2m}{n} + \sqrt{\frac{\log m \log(mn)}{n}} \log \frac{2m}{\log m} \right).$$

Thus, the same randomized polynomial-time algorithm finds a point $x \in \mathbb{Z}^n$ satisfying

$$|\alpha_i(x - x_0)| = \frac{|a_i(x - x_0)|}{\|a_i\|} \leq 2^{16} \left(\log \frac{2m}{n} + \sqrt{\frac{\log m \log(mn)}{n}} \log \frac{2m}{\log m} \right)$$

for every $i \in [m]$. The success probability of the algorithm reduces by a factor of $1 - 2me^{-n/96}$ due to the randomness in the input. \square

Acknowledgment. We are grateful to Shabbir Ahmed, Nikhil Bansal, Daniel Dadush, Santanu Dey and Joel Spencer for their kind help and encouragement.

References

- [1] N. Bansal. Constructive algorithms for discrepancy minimization. *Foundations of Computer Science (FOCS), 2010 51st Annual IEEE Symposium*, pages 3–10, 2010.
- [2] R. Beier and B. Vöcking. Random knapsack in expected polynomial time. In *STOC '03: Proceedings of the thirty-fifth annual ACM symposium on Theory of computing*, pages 232–241, New York, NY, USA, 2003. ACM.
- [3] R. Beier and B. Vöcking. Typical properties of winners and losers in discrete optimization. In *Proceedings of the thirty-sixth annual ACM symposium on Theory of computing*, STOC '04, pages 343–352, New York, NY, USA, 2004. ACM.

- [4] B. Bollobás. *Random graphs*. Cambridge studies in advanced mathematics. Cambridge University Press, 2001.
- [5] M. Charikar, A. Newman, and A. Nikolov. Tight hardness for minimizing discrepancy. In *Proceedings of the Twenty-Second Annual ACM-SIAM Symposium on Discrete Algorithms*, SODA '11, pages 1607–1614, Philadelphia, PA, USA, 2011. Society for Industrial and Applied Mathematics.
- [6] G. Dantzig. On the significance of solving some linear programs with some integer variables. *Econometrica*, 28:30–34, 1960.
- [7] S. Dasgupta and L. J. Schulman. A two-round variant of em for gaussian mixtures. In *Proceedings of the 16th Conference on Uncertainty in Artificial Intelligence*, UAI '00, pages 152–159, San Francisco, CA, USA, 2000. Morgan Kaufmann Publishers Inc.
- [8] M. L. Furst and R. Kannan. Succinct certificates for almost all subset sum problems. *SIAM J. Comput.*, 18:550–558, June 1989.
- [9] R. Gomory. An algorithm for integer solutions to linear programs. In *Recent advances in Mathematical Programming*, pages 269–302, New York, NY, USA, 1963. McGrawHill.
- [10] M. Grötschel, L. Lovász, and A. Schrijver. The ellipsoid method and its consequences in combinatorial optimization. *Combinatorica*, 1:169–197, 1981. 10.1007/BF02579273.
- [11] R. Kannan. Minkowski’s convex body theorem and integer programming. *Mathematics of Operations Research*, 12:415–440, 1987.
- [12] R. Karp. Reducibility among combinatorial problems. *Proc. Sympos. IBM Thomas J. Watson Research Center*, pages 85–103, 1972.
- [13] D. Kleitman. On a combinatorial conjecture of erdős. *Journal of Combin. Theory*, 1:209–214, 1966.
- [14] L. Lovász, J. Spencer, and K. Vesztegombi. Discrepancy of set-systems and matrices. *Eur. J. Comb.*, 7:151–160, April 1986.
- [15] J. Matoušek. An lp version of the beck-fiala conjecture. *Eur. J. Comb.*, 19:175–182, February 1998.
- [16] J. Matoušek and J. Spencer. Discrepancy in arithmetic progressions. *American Mathematical Society*, 9(1):195–204, January 1996.
- [17] G. Pataki, M. Tural, and E. B. Wong. Basis reduction and the complexity of branch-and-bound. In *Proceedings of the Twenty-First Annual ACM-SIAM Symposium on Discrete Algorithms*, SODA '10, pages 1254–1261, Philadelphia, PA, USA, 2010. Society for Industrial and Applied Mathematics.
- [18] H. Röglin and B. Vöcking. Smoothed Analysis of Integer Programming Integer Programming and Combinatorial Optimization. volume 3509 of *Lecture Notes in Computer Science*, chapter 21, pages 87–98. Springer Berlin / Heidelberg, Berlin, Heidelberg, 2005.

- [19] A. Schrijver. Theory of linear and integer programming. 1998.
- [20] J. Spencer. Six standard deviations suffice. *Trans. Amer. Math. Soc.*, 289:679–706, 1985.
- [21] J. Spencer. Ten lectures on the probabilistic method. *SBMS-NSF,SIAM*, 1987.